**Supplementary Material: On Multiset Selection with Size Constraints**

Chao Qian$^1$, Yibo Zhang$^1$, Ke Tang$^2$, Xin Yao$^2$

$^1$Anhui Province Key Lab of Big Data Analysis and Application, School of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China

$^2$Shenzhen Key Lab of Computational Intelligence, Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

chaoqian@ustc.edu.cn, zyb233@mail.ustc.edu.cn, {tangk3, xiny}@sustc.edu.cn

**Detailed Proofs**

This part aims to provide some detailed proofs, which are omitted in our original paper due to space limitation.

**Proof of Lemma 5.** By the definition of the DR-submodularity ratio (i.e., Definition 8 in the original paper), we have

$$
\beta_f = \min_{x \leq y \in [n]} \frac{f(x + \chi_i) - f(x)}{f(y + \chi_i) - f(y)}
$$

where

$$
\chi_i = \min_{x \leq y \in [n]} \frac{f(x + (y_i - x_i)\chi_i + \chi_i) - f(x + (y_i - x_i)\chi_i)}{f(x + \chi_i) - f(x)}.
$$

where the first inequality is by Lemma 2 since $f$ is submodular. We then calculate $f(x + m\chi_i) - f(x + (m-1)\chi_i)$ for any $1 \leq m \leq c_i - x_i$. By the definition of the objective function $f$ (i.e., Definition 4 in the original paper), we get

$$
f(x + m\chi_i) - f(x + (m-1)\chi_i) = \sum_{t \in T} \sum_{l} \lambda_{t,l}(x + m\chi_i) - \sum_{t \in T} \sum_{l} \lambda_{t,l}(x + (m-1)\chi_i),
$$

where the second equality is because for $x + m\chi_i$ and $x + (m-1)\chi_i$, only the probabilities of activating those target nodes adjacent to $v_i$ are different. We then calculate $f_t,i(x + m\chi_i) - f_t,i(x + (m-1)\chi_i)$ by the definition of $f_t,i(x)$ (i.e., Eq. (3) in the original paper). For notational convenience, we denote $x + m\chi_i$ and $x + (m-1)\chi_i$ by $y$ and $z$, respectively. Then, we have

$$
f_t,i(x + m\chi_i) - f_t,i(x + (m-1)\chi_i) = f_t,i(y) - f_t,i(z) = \prod_{r:(v_r,t) \in E} \prod_{j=1}^{m-1} (1 - p_r^{(j)}) \prod_{j=1}^{m} (1 - q_r^{(j)})
$$

$$
\geq \prod_{r:(v_r,t) \in E} \prod_{j=1}^{m-1} (1 - p_r^{(j)}) \prod_{j=1}^{m} (1 - q_r^{(j)})
$$

$$
\geq \prod_{r:(v_r,t) \in E} \prod_{j=1}^{m-1} (1 - p_r^{(j)}) \prod_{j=1}^{m} (1 - q_r^{(j)})
$$

Note that in the above equality, $x + m\chi_i$ and $x + (m-1)\chi_i$ are different only on the budget of the source node $v_i$, and thus only the probabilities of activating those target nodes adjacent to $v_i$ will be affected. By applying this equality to Eq. (1), we get

$$
\beta_f \geq \min_{x \leq y \in [n], 1 \leq m \leq c_i - x_i} \frac{p_i^{(x+1)}}{p_i^{(x+m)}} \prod_{j=1}^{m-1} (1 - p_i^{(j)}) \prod_{j=1}^{m} (1 - q_i^{(j)})
$$

$$
\geq \min_{x \leq y \in [n], 1 \leq m \leq c_i - x_i} \frac{p_i^{(x+1)}}{p_i^{(x+m)}} \prod_{j=1}^{m-1} (1 - p_i^{(j)}) \prod_{j=1}^{m} (1 - q_i^{(j)})
$$

$$
\geq \min_{x \leq y \in [n], 1 \leq m \leq c_i - x_i} \frac{p_i^{(x+1)}}{p_i^{(x+m)}} \prod_{j=1}^{m-1} (1 - p_i^{(j)}) \prod_{j=1}^{m} (1 - q_i^{(j)})
$$

**Proof of Lemma 6.** It is easy to see that the analysis on $\beta_f$ (i.e., Eq. (1)) in the proof of Lemma 5 still holds here, since it only relies on the submodularity of $f$. We then calculate

$$
f(x + m\chi_i) - f(x + (m-1)\chi_i) \text{ for any } 1 \leq m \leq c_i - x_i.
$$

By the definition of the objective function $f$ (i.e., Definition 5 in the original paper), we get

$$
f(x + m\chi_i) - f(x + (m-1)\chi_i) = \sum_{t \in T} \sum_{l} \lambda_{t,l}(x + m\chi_i) - \sum_{t \in T} \sum_{l} \lambda_{t,l}(x + (m-1)\chi_i)
$$

where the second equality is because for $x + m\chi_i$ and $x + (m-1)\chi_i$, only the probabilities of activating those target nodes adjacent to $v_i$ are different. We then calculate $f_t,i(x + m\chi_i) - f_t,i(x + (m-1)\chi_i)$ by the definition of $f_t,i(x)$ (i.e., Eq. (3) in the original paper). For notational convenience, we denote $x + m\chi_i$ and $x + (m-1)\chi_i$ by $y$ and $z$, respectively. Then, we have

$$
f_t,i(x + m\chi_i) - f_t,i(x + (m-1)\chi_i) = f_t,i(y) - f_t,i(z) = \prod_{r:(v_r,t) \in E} \prod_{j=1}^{m-1} (1 - p_r^{(j)}) \prod_{j=1}^{m} (1 - q_r^{(j)})
$$

$$
\geq \prod_{r:(v_r,t) \in E} \prod_{j=1}^{m-1} (1 - p_r^{(j)}) \prod_{j=1}^{m} (1 - q_r^{(j)})
$$

$$
\geq \prod_{r:(v_r,t) \in E} \prod_{j=1}^{m-1} (1 - p_r^{(j)}) \prod_{j=1}^{m} (1 - q_r^{(j)})
$$

$$
\geq \prod_{r:(v_r,t) \in E} \prod_{j=1}^{m-1} (1 - p_r^{(j)}) \prod_{j=1}^{m} (1 - q_r^{(j)})
$$

where for the third equality, $\xi_{i,m,l}$ is defined as

$$
\xi_{i,m,l} = \begin{cases} 
q_i^{(x+m)}, & x_i + m \geq l, \\
p_i^{(x+m)}, & \text{otherwise}.
\end{cases}
$$
and for the last equality, \( \eta_{i,m,l} \) is the product of \( m-1 \) terms in the form of either \( 1 - p_i^{(j)} \) or \( 1 - q_i^{(j)} \) (where \( x_i + 1 \leq j \leq x_i + m - 1 \)), since the only difference between \( z = x + (m - 1)x_i \) and \( x \) is the \( i \)-th entry, i.e., \( z_i = x_i + m - 1 \).

Let \( \delta_{i,t} = \lambda_{i,t} \prod_{j=1}^{m} (1 - p_i^{(j)}) \prod_{j=1}^{m} (1 - q_i^{(j)}) \). By applying the calculation result of \( f(x + m\chi_i) - f(x + (m-1)\chi_i) \) to Eq. (1), we can get

\[
\beta_f = \min_{x \in \mathbb{Z}_+^n, i \in [n], 1 \leq m \leq c_i - x_i} \sum_{t: (v_t, v_i) \in E} \sum_{l=1}^{t} \xi_{i,l,1}\delta_{i,t} = q_i^{(x_i+1)} \sum_{t: (v_t, v_i) \in E} \sum_{l=1}^{t} \delta_{i,t} + p_i^{(x_i+1)} \sum_{t: (v_t, v_i) \in E \setminus \{x_i\}} \sum_{l=1}^{t} \delta_{i,t}.
\]

The second equality is by \( \eta_{i,1,l} = 1 \) and \( \eta_{i,m,l} \leq 1 \) for \( m \geq 1 \). According to the definition of \( \xi_{i,m,l} \), we can divide \( \sum_{t: (v_t, v_i) \in E} \sum_{l=1}^{t} \xi_{i,1,l}\delta_{i,t} \) and \( \sum_{t: (v_t, v_i) \in E} \sum_{l=1}^{t} \xi_{i,m,l}\delta_{i,t} \) into three parts, respectively. That is,

\[
\sum_{t: (v_t, v_i) \in E} \sum_{l=1}^{t} \xi_{i,1,l}\delta_{i,t} = q_i^{(x_i+1)} \sum_{t: (v_t, v_i) \in E} \sum_{l=1}^{t} \delta_{i,t} + p_i^{(x_i+1)} \sum_{t: (v_t, v_i) \in E \setminus \{x_i\}} \sum_{l=1}^{t} \delta_{i,t};
\]

\[
\sum_{t: (v_t, v_i) \in E} \sum_{l=1}^{t} \xi_{i,m,l}\delta_{i,t} = q_i^{(x_i+m)} \sum_{t: (v_t, v_i) \in E} \sum_{l=1}^{t} \delta_{i,t} + p_i^{(x_i+m)} \sum_{t: (v_t, v_i) \in E \setminus \{x_i\}} \sum_{l=1}^{t} \delta_{i,t}.
\]

Note that the corresponding ratios of these three parts are \( q_i^{(x_i+1)} / q_i^{(x_i+m)} \), \( p_i^{(x_i+1)} / p_i^{(x_i+m)} \) and \( p_i^{(x_i+1)} / p_i^{(x_i+m)} \), respectively. Since their minimum must be not larger than the ratio of the sum of the three parts, we have

\[
\sum_{t: (v_t, v_i) \in E} \sum_{l=1}^{t} \xi_{i,1,l}\delta_{i,t} \geq \min \left\{ \frac{q_i^{(x_i+1)}}{q_i^{(x_i+m)}}, \frac{p_i^{(x_i+1)}}{q_i^{(x_i+m)}}, \frac{p_i^{(x_i+1)}}{p_i^{(x_i+m)}} \right\}
\]

\[
= \min \left\{ \frac{q_i^{(x_i+1)}}{q_i^{(x_i+m)}}, \frac{p_i^{(x_i+1)}}{p_i^{(x_i+m)}} \right\},
\]

where the equality is by the setting \( q_i^{(j)} \leq p_i^{(j)} \) of the problem in Definition 5. Thus, we get

\[
\beta_f = \min_{x \in \mathbb{Z}_+^n, i \in [n], 1 \leq m \leq c_i - x_i} \min \left\{ \frac{q_i^{(x_i+1)}}{q_i^{(x_i+m)}}, \frac{p_i^{(x_i+1)}}{q_i^{(x_i+m)}}, \frac{p_i^{(x_i+1)}}{p_i^{(x_i+m)}} \right\}
\]

For the generalized influence maximization problem as presented in Definition 1, we prove in Proposition 1 that the objective function \( f(x) = \mathbb{E}[\left\{ \bigcup_{t \geq 1} A(X_t) \right\}] \) is monotone submodular, if the fundamental propagation model Independence Cascade (Goldenberg, Libai, and Muller 2001) is used. For the Independence Cascade model as shown in Definition 2, it starts from a seed set \( A_0 = X \) and uses a set \( A_t \) to record the nodes activated at time \( t \); at time \( t+1 \), each inactive neighbor \( v_j \) of \( v_i \in A_t \) becomes active with probability \( p_{i,j} \); this process is repeated until no nodes get activated at some time.

**Definition 1 (Generalized Influence Maximization).** Given a directed graph \( G = (V, E) \), capacities \( c_i \) (\( i \in [n] \)), edge probabilities \( p_{i,j} \) for each \( (v_i, v_j) \in E \), and a budget \( k \), it is to find a multiset \( x \in \mathbb{Z}_+^n \) such that

\[
\arg \max_{x \in \mathbb{Z}_+^n} \mathbb{E}[\left\{ \bigcup_{t \geq 1} A(X_t) \right\}] \quad \text{s.t.} \quad |x| \leq k,
\]

where \( X_1 = \{v_i \mid x_i \geq 1\} \) and \( A(X_1) \) is the number of nodes activated by propagating from \( X_1 \).

**Definition 2 (Independence Cascade).** (Goldenberg, Libai, and Muller 2001) Given a directed graph \( G = (V, E) \) with edge probabilities \( p_{i,j} \) for each \( (v_i, v_j) \in E \) and a seed set \( X \subseteq V \), it propagates as follows:

1. let \( A_0 = X \) and \( t = 0 \).
2. repeat until \( A_t = \emptyset \)
3. for each edge \( (v_i, v_j) \) with \( v_i \in A_t \) and \( v_j \in V \setminus \bigcup_{r \leq t} A_r \)
4. \( v_j \) is added into \( A_{t+1} \) with probability \( p_{i,j} \).
5. let \( t = t + 1 \).

**Proposition 1.** If the Independence Cascade propagation model is used, the objective function \( f(x) = \mathbb{E}[\left\{ \bigcup_{t \geq 1} A(X_t) \right\}] \) of generalized influence maximization is monotone and submodular.

**Proof.** The monotonicity of \( f \) is trivial. We are to prove its submodularity. According to Lemma 2 in the original paper, we only need to prove that for any \( x \leq y \) and \( i \in [n] \) with \( x_i = y_i \),

\[
f(x + \chi_i) - f(x) \geq f(y + \chi_i) - f(y).
\]

Let \( X_i = \{v_j \mid x_i \geq 1\} \). According to the definition of the objective function \( f \), we get

\[
f(x + \chi_i) - f(x) = \mathbb{E}[\left\{ A(X_{x_i+1} \cup \{v_i\}) \right\}] - \mathbb{E}[\left\{ A(X_{x_i+1} \cup \{v_i\} \mid x_i \neq 1\} \right\}] - \mathbb{E}[\left\{ A(X_{x_i+1} \cup \{v_i\} \mid x_i = 1\} \right\}].
\]

Let \( G' = (V, E') \) denote a subgraph of \( G = (V, E) \), which is generated by preserving each edge \( (v_i, v_j) \in E \) with probability \( p_{i,j} \). Then, the set of nodes reachable from \( X_t \) in \( G' \) actually corresponds to \( A(X_t) \). Note that \( G' \) is random. For each fixed \( G' \), it is easy to see that \( A(X_{x_i+1} \cup \{v_i\}) \subseteq A(X_{x_i+1} \cup \{v_i\}) \), given \( X_{x_i+1} \subseteq X_{x_i+1} \cup \{v_i\} \). Let \( S = \bigcup_{i \neq 1} A(X_i) \). Thus, we have

\[
f(x + \chi_i) - f(x) = \mathbb{E}[\left\{ A(X_{x_i+1} \cup \{v_i\}) \right\}] - \mathbb{E}[\left\{ A(X_{x_i+1} \cup \{v_i\} \mid S \right\}].
\]
Let $Y_l = \{ v_j \mid y_j \geq l \}$ and $T = \bigcup_{l \neq y_i + 1} A(Y_l)$. We can similarly get

$$f(y + \chi_i) - f(y) = \mathbb{E}\left[ |A(Y_{y_i + 1} \cup \{v_i\}) \setminus A(Y_{y_i + 1}) \setminus T| \right].$$

Since $x_i = y_i$, $T$ is actually $\bigcup_{l \neq x_i + 1} A(Y_l)$, and

$$f(y + \chi_i) - f(y) = \mathbb{E}\left[ |A(Y_{x_i + 1} \cup \{v_i\}) \setminus A(Y_{x_i + 1}) \setminus T| \right].$$

Note that $X_l \subseteq Y_l$, since $x \leq y$. Thus, for any $l$, it holds that $A(X_l) \subseteq A(Y_l)$ for each fixed subgraph. This implies that $S \subseteq T$. Furthermore, for each fixed subgraph, $A(Y_{x_i + 1} \cup \{v_i\}) \setminus A(Y_{x_i + 1}) \subseteq A(X_{x_i + 1} \cup \{v_i\}) \setminus A(X_{x_i + 1})$, since $A(Y_{x_i + 1}) \subseteq A(Y_{x_i + 1})$. Thus, we can get

$$(f(x + \chi_i) - f(x)) - (f(y + \chi_i) - f(y)) \geq 0.$$ 

Thus, the proposition holds. 

References